DOCUMENT RESUME

ED 206 165 FL 012 450

AUTHOR Genesee, Fred
TITLE Psycholinguistic Foundations of Language Assessment.
PUB DATE Jun 81
NOTE 21p.: Paper presented at the Language Assessment Institute, National College of Education (Evanston, IL, June 17, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Applied Linguistics; *Communicative Competence (Languages); Elementary Secondary Education; Language Research; *Language Tests; Postsecondary Education; *Psycholinguistics; Second Language Learning; Sociolinguistics; *Testing

ABSTRACT

A review of literature on foreign language testing indicates that the earliest approaches to language assessment were generally uninformed by contemporaneous linguistic and psychological theories and were characterized by a lack of psychometric sophistication. This trend was followed by the development of test instruments that were heavily influenced by structural linguistics and behavioral psychology. Accordingly, such tests contained multiple-choice items pertaining to discrete linguistic structures and skills. This type of test prevailed for some time and is still widely used. The theoretical basis argued that they failed to reflect the integrative, communicative aspects of real language use. The early criticism of the sociolinguists was reinforced by that of theoretical linguists, psycholinguists, and educators who are adopting communicative approaches to the study of language within their respective domains. The field of communicative competence testing is relatively embryonic and faces formidable challenges to its development, the principal one being the lack of a widely accepted model of communicative competence. Notwithstanding theoretical and psychometric problems, the notion of communicative competence provides a necessary next step in the evolution of language assessment techniques. (Author/AMH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Psycholinguistic Foundations of Language Assessment

Fred Genesee

McGill University

ED206165

Whether carried out informally by classroom teachers using tests of their own construction or formally by trained evaluators using standardized tests, testing has become an important and routine part of the modern educational process. Rebecca Valette in her 1967 book entitled Modern Language Testing identified three important roles of classroom tests in second language programs. First, language tests define course objectives by indicating to students the criterion, or criteria, on which they will be evaluated and, thereby, passed or failed. For example, second language programs which utilize tests of reading and writing proficiency to assess student achievement will direct the students to study these aspects of the second language even, as Valette points out, in cases where the ostensible goals of the program may be defined in other terms, such as oral proficiency. Second, classroom tests stimulate student progress by providing feedback concerning their mastery or lack of mastery of particular elements of the language under study. Third, classroom tests evaluate language achievement and in so doing indicate to the instructor which elements of the program are being learned successfully by the class as a whole and/or by individual students. Provided with this information, the instructor is then better able to tailor his/her teaching to meet the needs of the students. In other words,

FL012450

8-21-29

Paper presented at the Language Assessment Institute, National College of Education, Evanston, Illinois, June 17, 1981.

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Fred Genesee

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

tests can provide feedback to the teacher as well as to the student.

Language testing serves an additional role, one that is of particular interest to those involved in the education of minority language children. Tests can help identify the educational needs of individual or subgroups of children and, thereby, inform decision-makers regarding the appropriateness of alternative program options. For example, entry to or exit from Title VII Bilingual programs depends in many cases on the students' performance on language tests. Similar placement testing is of course characteristic of decision-making in the case of mainstream children as well--decisions whether a student will receive enriched or remedial instruction, for example. Because of the possible long term educational and even social implications some of these decisions entail, it is important to know and understand the linguistic and psychological origins of different approaches to language assessment.

Although the theme of this lecture is the psycholinguistic foundations of language assessment, it is interesting to note at the outset, and this will become evident shortly, that the field of language testing has been influenced primarily and directly by linguistics. The primary role that linguistic theories have played in the development of language tests is rather paradoxical when one considers that traditionally linguists have not concerned themselves with the implications of their theories for actual language use, or performance in the

Chomskyian sense, and, in fact, they have openly disavowed any responsibility for pragmatic issues. The conventional task of the linguist has been to describe the abstract grammatical system that the idealized speaker-listener is supposed to possess. Linguistic evidence for or against particular grammatical rules consisted of the linguists' own informed intuitions. The psychological reality of these rules was of no concern. Issues pertaining to actual language use, or performance, fall naturally within the psycholinguist's domain, and yet work in this field has had less substantive effect on language testing than has linguistics. In fact, with the exception perhaps of discrete-point tests, to be described shortly, it could be said that language tests are not based on any coherent or explicit psycholinguistic theory (Ingram, 1978). Psychological and psycholinguistic work has, however, provided theoretical support for the types of testing procedures promulgated by linguistic theories. Psychology has also had a direct impact on language testing through the influence of psychometrics.

The earliest approach to language assessment has been described by Spolsky (1978) as prescientific. This approach was characterized by a lack of concern for statistical validity or reliability and, generally speaking, was uninformed by linguistic or psycholinguistic theories of language. In particular, according to Spolsky, it assumed "that one can and must rely completely on the judgment of an experienced teacher, who can

tell after a few minutes' conversation, or after reading a student's essay, what mark to give" (p.v). Thus, this approach did not require the services of expert test developers or administrators, but rather fell clearly within the competence of any and all capable classroom teachers. Language assessment of this type continues to be practiced in certain areas of the world and also by individual teachers in educational settings where more "scientific" approaches predominate. It is interesting to note that the lack of formal statistical bases of these tests and their corresponding reliance on teacher's subjective judgments, which have been sources of serious criticism, are presently viewed less critically by educators who support the use of informal language assessment for classroom purposes.

This type of language assessment gave way to tests that were influenced substantially by formal linguistic and psychological theories, and, in particular, by the American structuralist school of linguistics (Bloomfield, 1933) and the behaviorist school of psychology (Watson, 1913). The influence of these schools of thought on language testing was given its most complete expression in the work of Robert Lado which was published in 1961 in his now famous book on Language Testing. In accord with structural linguistics and behavioural psychology, Lado contended that language is made up of a number of discrete structures that are arranged in an hierarchical fashion. He differentiated two principal units of structural analysis which he referred to as elements and skills. The

main elements of language included pronunciation, grammar, the lexicon, and what he referred to as cultural meaning. Each of these elements in turn could be analyzed into smaller elements. For example, pronunciation is made up of sound segments (or phonemes), intonation and stress; grammar is made up of morphology and syntax. Lado felt that "The elements of language...can be profitably studied and described and tested as separate universes" (p.25). At the same time, he felt that the elements never actually occur separately in language but rather are integrated to constitute the skills of speaking, listening, reading and writing. An additional skill that Lado differentiated was translation. All five of these skills could and should be the object of language assessment according to Lado.

Lado's structural analysis of language provides a matrix of "variables" (p. 25), consisting of language skill-by-element combinations, each of which could be tested separately; not all elements can necessarily be tested in all skill areas.

ELEMENT	SKILL: Reading	Writing	Speaking	Listening
sounds				
intonation				
stress				
morphemes				
words				
word order				

One additional feature of both the structuralist and behaviorist approaches to language had a significant influence on the development of language tests at this time. Both schools of thought claimed that "language is a system of habits of communication" (Lado, p.22) and that these habits could be described and explained by studying samples of speech or authentic utterances "gathered in an objective manner without historical or theoretical biases" (Palermo, 1978, p.21). More specifically, the analysis of linguistic data was supposed to proceed without regard to mentalistic, i.e., cognitive, social or other psychological factors. Thus, language was regarded as a self-contained system of linguistic structures and rules and language testing, therefore, could and should be accomplished in isolation of non-linguistic factors.

These various claims regarding the nature of language led naturally to the formulation of language tests of a particular type. Thus, we find that tests developed from this perspective consist of a series of subtests each measuring ostensibly separate skills and consisting of a series of discrete items. The bias toward evaluation of discrete linguistic elements was reinforced by psychometric concerns regarding the reliability and validity of previous testing techniques, which, as we have already seen, relied on individual teacher judgments and intuitions. It was much easier to devise reliable and valid tests, according to the psychometricians, if they consisted of discrete items that could be scored objectively. As a result,

multiple-choice questions were promoted as the most appropriate means to assess language, and they have become the hallmark of this tradition. Discrete-point multiple-choice tests have the added advantage of being easy to administer and score, and they permit group testing. Spolsky has charaterized this approach to test development as the structuralist-psychometric trend.

Although discrete-point tests of the type just described were well received from the outset and indeed have served to advance the field of language assessment from a prescientific to a scientific paradigm, there has also been criticism of this approach from the beginning (J.B. Carroll, 1961), and it has grown considerably in recent years. Initial criticisms were directed at the notion that language is constituted of discrete, separable structures and that language proficiency, therefore, can be defined in terms of knowledge of these discrete items. In fact, there is little empirical evidence that the components of language that were described by Lado and the structural linguists are actually distinct (Oller, 1977). Regardless of the psychological distinctiveness of these putative linguistic elements, it is argued that functional language use entails the complex integration of linguistic rules in conjunction with higher order rule systems, including those related to interpersonal communication, or discourse, sociolinguistic characteristics of the situation and of the interactants, and other nonlinguistic, or pragmatic features

of the interaction. Tests that are restricted to sentence level processing of language as is the tendency with tests of the structural linguistic persuasion, are limited to assessing low level linguistic competence and ignore the examinees' ability to use linguistic rules effectively in the service of real communication. In short, previous testing techniques failed to assess communicative competence.

An early response to these criticisms, and particularly that one pertaining to the inadequacy of assessing discrete language elements and skills, has been the development of so-called integrative language tests. The two most familiar ones being the cloze test (Oller, 1973) and the dictation tests which may be administered with (Spolsky, Sigurd, Sato, Walker, & Arterburn, 1968) or without distracting noise (Oller, 1973). By using discourse length material these tests purport to engage more integrative and, therefore, more communicative language skills. While it may be true that these tests do assess integrative use of language in some sense, they are still grossly inadequate to tap communicative competence of the sort that was alluded to by critics of the discrete-point tradition.

The notion of communicative competence is not new. Sociolinguists have long recognized the fact that speakers have a number of speech styles or language varieties at their disposal and that communicative competence is characterized by dif-

ferential and systematic use of these styles depending upon the topic and purpose of conversation, the situation in which communication takes place and on characteristics of the interlocutors (Fishman, 1972; Hymes, 1972). The work of theoretical linguists, such as Grice (1975) and Searle (1969) has also emphasized the importance of the interaction between linguistic, discourse, social and psychological variables in interpersonal communication. Psycholinguists, as well, have recently turned their attention to communicative aspects of language acquisition in contrast to the strictly linguistic focus of earlier research (Ochs & Scheffelin, 1978). Certainly the notion of communicative competence has been pervasive in the field of second language education. The pioneering work of Sandra Savignon (1972) and the Canadian experiments in second language immersion (Lambert & Tucker, 1972) attest to the longstanding influence that communicative approaches have had on classroom teaching.

In contrast, it is only recently that serious thinking on a large scale has been directed toward the assessment of communicative competence. The state of communicative competence testing can be best characterized perhaps by B.J. Carroll's comment that "It is unfortunate that the adventurous, thinking about communicative competence that is taking place in the field of applied linguistics, communication and sociolinguistics is not being matched in the field of testing" (1980, p.90). The development of sound communicative tests

faces a number of formidable conceptual and psychometric challenges.

Despite the considerable theoretical attention that has been devoted to models of communicative competence from diverse scholarly disciplines, as referred to earlier, there is still considerable debate over the precise form that such a model should take (see Canale & Swain, 1980, and Wesche, 1981, for recent discussions of this topic). In particular, there is as yet no consensus concerning the types of competencies that should be included in such a model. Some theorists, such as Hymes (1972), include linguistic competence, that is knowledge of the structure of a given language in the structural linguistic tradition, along with sociolinguistic competence, that is knowledge underlying the effective and appropriate use of linguistic rules in social situations. Canale and Swain (1980) have argued in favour of also including strategic competence, that is verbal and non-verbal communication strategies that may be used to compensate for lack of competence in linguistic rules. Others accept a definition of communicative competence that includes the ability to transmit messages successfully using any means, linguistic or non-linguistic, without regard to their grammaticality or sociolinguistic appropriateness (Savignon, 1972; Schultz, 1977). Yet others maintain that the distinction between linguistic competence and sociolinguistic competence is ill-founded. Connors (1980) contends that the

two are linguistically indissociable and that evidence of distinctiveness actually reflects differences between language and non-language based proficiencies. Without delving further into this thorny theoretical issue, which seems far from being resolved, the point should be fairly clear--no comprehensive model of communicative competence assessment can be developed until these issues are settled.

In lieu of a definitive model of communicative competence, we might accept an interim model that includes the minimal distinction between linguistic and sociolinguistic competence. Even in this case two fundamental theoretical issues remain. The first concerns the identification and description of the language functions that constitute basic communicative competence. Language functions include such communicative acts as apologizing, requesting, describing, and they constitute the basic unit of communicative theories of language proficiency; this contrasts with structuralist theories which are based on phonological, grammatical, lexical and semantic elements. Wilkins (1976, p.23) points out that "while there has been quite enough scholarly discussion of semantic and grammatical categories for there to be little that is contentious in a list of candidates for inclusion in a semantic-grammatical inventory, there is no comparable authority for establishing a universal set of categories of language use". Lists of language functions have been prepared for pedagogical purposes by van Ek (1976) and Wilkins,

but it must be admitted that these are in fact ad hoc. There is as yet no empirical or theoretical justification for the functions in these lists.

The second theoretical issue concerns the specification of the linguistic structures that are necessary to perform designated language functions in a sociolinguistically appropriate fashion. In other words, we lack a grammar of communication. Connors (1980) believes that it is probably impossible to formulate such a grammar in the present state of our knowledge of communicative competence and in view of the intuitive way in which language functions have been formulated. Extant grammars are likely to offer little assistance since they have been formulated on intuitive, asocial grounds. The best we may be able to achieve at this time is an empirical description of the typical linguistic structures that a sample of native speakers use to realize particular functions in certain social situations. How we describe in a concise and heuristic fashion the relationship between the structures thus identified and their respective functions is an open question. The importance for language assessment of these lacunae in our understanding of communicative competence is that we cannot test what we cannot describe.

The psychometric problems facing test developers are no less formidable than the theoretical problems just reviewed. At least three major ones come to mind: (1) the authenticity

or face validity of communicative competence tests; (2) their representativeness or situational generalizability; and (3) the definition of proficiency levels.

Authenticity. Although there is as yet little consensus regarding the precise nature of communicative competence, it is nevertheless generally accepted that real communication is dynamic, interactive and often even unpredictable. The challenge facing psychometricians is how to incorporate these features of communication into a test. It has been suggested that the development of direct communicative language tests would circumvent the problem of artificiality created by indirect tests, such as the cloze or dictation test (B.J. Carroll, 1980; Clark, 1972). While it is true that direct tests would have greater face validity, they may be impractical or unfeasible given the constraints operating in most testing settings. Furthermore, in contrast to most indirect tests that can be relatively easily scored by untrained personnel using objective criteria, the scoring of direct tests often requires the services of highly trained persons. In order to be useful and effective a test must be practical.

Situational generalizability. An additional aspect of real communication that must be considered by psychometricians is its dependency on characteristics of the situation and of others involved in the communication. The characteristics of appropriate language use change as the setting, topic of con-

versation and interlocutors change. The precise way in which these changes co-occur is an empirical question, still to be investigated. A comprehensive test of communicative competence must reflect this situational dependency. Therefore, it is unlikely that an assessment of communicative competence will consist of one test; rather it is likely to require many, each reflecting sociolinguistically distinct situations. The necessity of using multiple tests to assess communicative competence would strain practicality even further.

Definition of proficiency levels. Assessment of communicative competence differs from traditional language proficiency assessment in that it is not concerned so much with correctness as with appropriateness--that is, what set of linguistic structures are necessary or sufficient to realize particular language functions in sociolinguistically appropriate ways. Because of the lack of a definitive model of communicative competence, we lack a systematic conceptual basis for establishing proficiency levels that are criterion referenced. Lacking theoretical criteria, we could nevertheless establish norms of proficiency by examining empirically the types of linguistic structures that native speakers use to realize particular language functions in certain social situations. Proficiency levels might be defined according to the relative frequency of occurrence of specific structures or methods of expression of different functions. Even this empirical approach, which

seems practical, albeit demanding, and conceptually acceptable is problematic. When choosing a norming group, should one select a random group of native speakers, or should they be selected according to certain characteristics, such as their academic ability, socioeconomic background or even sex. In the case of communicative competence in a second language are native speakers in fact the appropriate norming group? Canale and Swain (1980) point out that native-speaking interlocutors may be more tolerant toward linguistic and sociolinguistic "failures" in the communications of second language speakers than in that of native language speakers. Thus, the use of native speakers as a norming group may constitute an unrealistically high and inappropriate standard for assessing second language speakers.

My purpose in reviewing in some length the problems faced by communicative competence test developers is not to discourage enthusiasm for this approach, but rather to highlight those characteristics that set it apart from more traditional and currently more prevalent techniques. It is evident that if language assessment is to keep stride with evolving perspectives in linguistics, sociolinguistics and psychology, it will necessarily have to take into account the notion of communicative competence, however we finally define it. One should not be surprised to find apparently insurmountable problems obstructing our progress in this field since real

communication, as we all know, involves no less than all those factors associated with individuals engaged in diverse interpersonal encounters with varied social characteristics. Indeed, the very complex and problematic nature of this approach suggests that we are in fact moving in the right direction.

Summary

Early, so-called pre-scientific approaches to language assessment were generally uninformed by contemporaneous linguistic and psychological theories and were characterized by a lack of psychometric sophistication. This trend was followed by the development of test instruments that were heavily influenced by structural linguistics and behavioural psychology. Accordingly, such tests contained multiple-choice items pertaining to discrete linguistic structures and skills. This trend has prevailed and is currently prevalent still. The theoretical basis of these tests were criticized primarily by sociolinguists who argued that they failed to reflect the integrative, communicative aspects of real language use. The early criticisms of the sociolinguists have been reinforced by theoretical linguists, psycholinguists and educators all of whom are similarly adopting communicative approaches to the study of language within their respective perspectives. It was pointed out that the field of communicative competence testing is relatively embryonic and faces formidable challenges to its development, the principal one being the lack of a

widely accepted model of communicative competence. Notwithstanding theoretical and psychometric problems, the notion of communicative competence provides a necessary next step in the evolution of language assessment techniques.

References

Bloomfield, L. Language. New York: Holt, Rinehart & Winston, 1933.

Canale, M., & Swain, M. Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1980, 1, 1-47.

Carroll, B.J. Testing communicative performance: An interim study. Oxford: Pergamon Press, 1980.

Carroll, J.B. Fundamental considerations in testing for English language proficiency of foreign students. Testing, Washington, D.C.: Center for Applied Linguistics, 1961. Reprinted in H.B. Allen & R.N. Campbell (Eds.), Teaching English as a second language: A book of readings, 2nd ed. N.Y.: McGraw Hill, 1972.

Clark, C.L.D. Foreign language testing: Theory and practice. Philadelphia: The Center for Curriculum Development, 1972.

Connors, K. Grammatical vs communicative competence in second language learning: A view from linguistics. Canadian Journal of Psychology, 1980, 34, 328-336.

Fishman, J.A. The sociology of language. In P.P. Giglioli (Ed.), Language and social context. Middlesex: Harmondsworth, 1972.

Grice, H.P. Logic and conversation. In P. Cole & J.J. Morgan (Eds.), Syntax and semantics, Vol. 3: Speech arts. N.Y.: Academic Press, 1975.

Hymes, D. On communicative competence. In J.B. Pride & J. Holmes (Eds.), Sociolinguistics. Harmondsworth, Eng.: Pergamon Books, 1972.

Ingram, E. The psycholinguistic basis. In B. Spolsky (Ed.), Advances in language testing. Series 2. Arlington, Virginia: Center for Applied Linguistics, 1978, 1-14.

Lado, R. Language testing. London: Longmans, Green, 1961.

Lambert, W.E., & Tucker, G.R. The bilingual education of children: The St. Lambert experiment. Rowley, Mass.: Newbury Hourse, 1972.

Ochs, E., & B. Scheffelin (Eds.), Developmental pramatics. N.Y.: Academic Press, 1979.

Oller, J.W. Cloze tests and second language proficiency and what they measure. Language Learning, 1973, 23, 105-118.

Oller, J.W. Language testing. In R. Wardhaugh & H.D. Brown (Eds.), A survey of applied linguistics. Ann Arbor: University of Michigan Press, 1977, 275-300.

Palermo, D.S. Psychology of language. Dallas, TX: Scott, Foresman, 1978.

Savignon, S. Communicative competence: An experiment in foreign-language teaching. Philadelphia: The Center for Curriculum Development, 1972.

Schultz, R.A. Discrete-point versus simulated communication testing in foreign languages. The Modern Language Journal, 1977, 61, 94-100.

Searle, J.R. Speech acts. Cambridge: Cambridge University Press, 1969.

Spolsky, B. (Ed.), Advances in language testing, Series 2. Arlington, Va.: Center for Applied Linguistics, 1978, v-$\bar{x}$.

Spolsky, B., Sigurd, B., Sato, M., Walker, E., & Arterburn, C. Preliminary studies in the development of techniques for testing overall second language proficiency. In J.A. Upshur & J. Fata (Eds.), Problems in foreign language testing language learning. Special issue No. 3, 1968, 79-101.

Valette, R.M. Modern language testing. New York: Harcourt, Brace, Javanovitch, 1967.

van Ek, J. The threshold level for modern language learning in schools. Groningen: Longman, 1977.

Watson, J.B. Psychology as the behaviorist views it. Psychological Review, 1913, 20, 158-177.

Wesche, M.B. Communicative testing in a second language. The Canadian Modern Language Review, 1981, 37, 551-571.

Wilkins, D.A. Notional syllabuses. Oxford: Oxford University Press, 1976.